Nonlinear Hilbert adjoints

Scherpen, Jacquelien M.A.; Gray, W. Steven

Published in:
Nonlinear Analysis

DOI:
10.1016/S0362-546X(01)00867-7

PERGAMON

Nonlinear Analysis 51 (2002) 883 – 901

Nonlinear Analysis

www.elsevier.com/locate/na

Nonlinear Hilbert adjoints: properties and applications to Hankel singular value analysis

Jacquelien M.A. Scherpen[a,*], W. Steven Gray[b]

[a] *Delft University of Technology, School of Electrical Engineering, Faculty of Information Technology and System, P.O. Box 5031, 2600 GA Delft, The Netherlands*

[b] *Department of Electrical, and Computer Engineering, Old Dominion University, Norfolk, VA 23529-0246, USA*

Received 7 February 2001; accepted 4 October 2001

Keywords: Adjoint maps; Hankel operators; Singular value functions; Nonlinear systems

1. Introduction

The notion of an *adjoint map* or *adjoint operator* can be found in a wide variety of mathematical contexts: functional analysis [39], differential geometry [1], differential algebra [42], representation theory for Lie algebras [5] and topological vector spaces [32]. These concepts appear primarily in a linear setting, i.e., linear maps on linear spaces, and thus are closely related to one another. In linear system theory, the important notion of an adjoint *state-space system* is usually defined in terms of signal sets that form Hilbert spaces, either L_2 or H_2 [43]. From an input–output point of view, the corresponding transfer function follows directly from the familiar Hilbert adjoint in functional analysis.

A particularly important operator in a system-theoretic setting is the Hankel operator. In the theory of continuous-time linear systems, the system Hankel operator plays a central role in a number of realization problems. The compact Hankel operator supplies a set of similarity invariants, the Hankel singular values, which can be used to quantify the importance of each state in the corresponding input–output map [25]. They determine the minimal dimension of any corresponding state-space system, and provide, via so-called balanced realizations, [25,28], a useful tool for model reduction of the

* Corresponding author. Tel.: +31-15-278-6152; fax: +31-15-278-6679.

E-mail addresses: j.m.a.scherpen@its.tudelft.nl (J.M.A. Scherpen), gray@ece.odu.edu (W.S. Gray).

PII: S0362-546X(01)00867-7

linear system. The Hankel singular values can be computed in a state-space setting using the product of the controllability and observability Gramian matrices, though intrinsically they depend only on the Hankel operator and its adjoint.

Once one departs from the context of linear operators, there are some extensions of the adjoint operator definition. It cannot be assumed a priori that the existing notions are in any way directly related. For example, in [4] the notion of an adjoint map is defined in terms of a dual map on a topological vector space. This idea is distinct from the adjoint map that appears in [7] which employs the Gâteaux derivative of the operator when it is well defined. Other distinct definitions can be found in [2,9]. A set of definitions that is useful for system theoretical considerations, and in particular realization theory, is given in [8,27,29,38,40], although these papers are not addressing this application. In a nonlinear state-space context, the adjoint system has appeared in [10], but only recently has it been given an input–output interpretation using the nonlinear Hilbert adjoint operator [13–16]. This latter concept first appeared in an abstract setting in [19,22,35] mainly to address the open problem of understanding how to relate the state-space notion of singular value functions due to Scherpen [33] to the nonlinear Hankel operator extension. But a broader investigation of this concept was not pursued. So in this paper, the basic objective is to fully develop the idea of a nonlinear Hilbert adjoint and to further illustrate its usefulness in Hankel singular value analysis.

The paper is structured as follows: in Section 2, the existing background for the paper is briefly summarized. This includes the definition of the nonlinear Hilbert adjoint, a local existence theorem, and its application to singular value analysis of the nonlinear Hankel operators. Section 3 is devoted entirely to describing the basic properties of the adjoint operator and presenting some simple examples. Section 4 then uses some of these new properties to further extend the existing Hankel operator analysis. Section 5 summarizes the conclusions of the paper.

The mathematical notation used throughout is fairly standard. Vector norms are represented by $\|x\| = \sqrt{x^T x}$ for $x \in \mathbb{R}^n$. $L_2^m(a,b)$ represents the set of m vector-valued Lebesgue measurable functions with finite norm $\|x\|_{L_2^m} = \sqrt{\int_a^b \|x(t)\|^2 \, \mathrm{d}t}$. (The superscript is suppressed when $m = 1$.) If $L : \mathbb{R}^n \mapsto \mathbb{R}$ is a differentiable function, then its partial derivative $\partial L/\partial x$ will be the row vector of partial derivatives $\partial L/\partial x_i$ where $i = 1, \ldots, n$. More generally, if L is a mapping between two Banach spaces, then its Fréchet derivative at a point u is denoted by $\mathrm{D}G(u)$.

2. Definitions and background material

In this section some established background material concerning nonlinear Hilbert adjoint operators and nonlinear Hankel operators is briefly reviewed.

Defining the nonlinear Hilbert adjoint: In the most general setting, let F be a topological vector space over $\mathbb{R}$ with dual space F' [32]. Let E be a nonempty set, and $\mathscr{A}$ a collection of nonempty subsets of E. Let E^β be a linear space of real-valued functions x^β on E with the property that the restriction x_A^β to every $A \in \mathscr{A}$ is bounded.

A mapping $\mathscr{T}: E \to F$ is called $\mathscr{A}$*-bounded* if $\mathscr{T}$ maps the sets of $\mathscr{A}$ into bounded subsets of F. For any $\mathscr{A}$-bounded mapping $\mathscr{T}: E \to F$, the *dual map* of $\mathscr{T}$ is defined as

$$\mathscr{T}': F' \to E^{\beta}$$

$$: y' \to (\mathscr{T}'(y'))(u) = (y' \circ \mathscr{T})(u), \quad \forall u \in E$$

(see, for example, [4]). Now if F is a Hilbert space with an inner product $\langle \cdot, \cdot \rangle_F$ then it follows from the Riesz Lemma that for any $y' \in F'$ there exists a unique $y \in F$ such that $y'(\cdot) = \langle y, \cdot \rangle_F$. Hence one can write the identity

$$(\mathscr{T}'(y'))(u) = \langle y, \mathscr{T}(u) \rangle_F, \quad \forall u \in E.$$

If, in addition, E is an inner product space with inner product $\langle \cdot, \cdot \rangle_E$ and $y \in F$ is fixed, then the problem is to determine a corresponding $\tilde{u}_y \in E$ such that

$$\langle \mathscr{T}(u), y \rangle_F = \langle u, \tilde{u}_y \rangle_E, \quad \forall u \in E. \tag{1}$$

If $\mathscr{T}$ were a linear operator then such an $\tilde{u}_y$ is known to always exist and be unique, i.e., $\tilde{u}_y = \mathscr{T}^*(y)$, where $\mathscr{T}^*$ is the Hilbert adjoint of $\mathscr{T}$. But in this more general context, the existence and uniqueness of $\tilde{u}_y$ are not automatic. In fact, identity (1) is meaningful in most cases only when $\tilde{u}_y$ is also a function of u. (Defining the domain of $\mathscr{T}^*$ to have the form $F \times E$ also agrees with the state-space notion of adjoint *systems* based on the Hamiltonian extension given in [10,37].) So in this context, consider the following definition.

Definition 2.1. Given two Hilbert spaces E and F, an operator $\mathscr{T}: E \mapsto F$ has a global nonlinear Hilbert adjoint when there exists an operator $\mathscr{T}^*: F \times E \to E$ such that

$$\langle \mathscr{T}(u), y \rangle_F = \langle u, \mathscr{T}^*(y, u) \rangle_E, \quad \forall u \in E, \quad \forall y \in F, \tag{2}$$

where $\mathscr{T}^*(y, u)$ is linear in y.

The above definition is more general than the definition of an adjoint operator given in [7], where identity (2) is only required to hold when $y = u$. To study singular value structures, $y = \mathscr{T}(u)$ should also be admissible. The adjoint definition of [7] is too limited for this purpose. Definition 2.1 is slightly different from the definition that appeared in [19,22,35] since here linearity in y is an additional requirement. But it seems rather natural in light of the bi-linearity of inner products, i.e.,

$$\begin{aligned}\langle u, \mathscr{T}^*(\alpha_1 y_1 + \alpha_2 y_2, u) \rangle_E &= \langle \mathscr{T}(u), \alpha_1 y_1 + \alpha_2 y_2 \rangle_F \\ &= \alpha_1 \langle \mathscr{T}(u), y_1 \rangle_F + \alpha_2 \langle \mathscr{T}(u), y_2 \rangle_F \\ &= \langle u, \alpha_1 \mathscr{T}^*(y_1, u) + \alpha_2 \mathscr{T}^*(y_2, u) \rangle_E.\end{aligned}$$

Linearity in y, however, does not follow directly from this argument. This is because there often exists a collection of nontrivial mappings (linear and nonlinear in y) of the form $\mathscr{B}: F \times E \mapsto E$ such that $\langle u, \mathscr{B}(y, u) \rangle_E = 0$, $\forall u \in E$, $\forall y \in F$. In which case, any adjoint mapping $\mathscr{T}^*$ is not uniquely defined since $\mathscr{T}^* + \mathscr{B}$ will also satisfy

Eq. (2). In these circumstances, an adjoint operator should be viewed as a member of an equivalence class, where two such operators $\mathscr{T}^*$ and $\mathscr{T}^{*'}$ are equivalent if

$$\langle u, \mathscr{T}^*(y,u)\rangle_E = \langle u, \mathscr{T}^{*'}(y,u)\rangle_E, \quad \forall u \in E, \quad \forall y \in F. \tag{3}$$

A shorthand notation for (3) is simply $\mathscr{T}^*(y,u) \cong \mathscr{T}^{*'}(y,u)$. Thus, any equality involving adjoint operators really means that both expressions belong to the same equivalence class. (See [20,21] for analysis and examples closely related to this issue.) The following example demonstrates this phenomenon.

Example 2.1. Consider the operator

$$\begin{aligned}\mathscr{B}: L_2[0,T] \times L_2^2[0,T] &\mapsto L_2^2[0,T]\\ :(y,u) &\mapsto c(u)A(u)b(y),\end{aligned}$$

where

$$A(u) = \begin{bmatrix} -u_2 & u_1u_2 \\ u_1 & -u_1^2 \end{bmatrix}$$

and c and b are any suitable mappings on $L_2^2[0,T]$ and $L_2[0,T]$, respectively. It can be verified directly that

$$\langle u, \mathscr{B}(y,u)\rangle_{L_2^2} = 0, \quad \forall u \in L_2^2[0,T], \quad \forall y \in L_2[0,T]$$

and thus any $\mathscr{T}^*(y,u)+\mathscr{B}(y,u)$ fulfills (2) when $\mathscr{T}^*(y,u)$ fulfills (2), even if $b(y)$ is a nonlinear mapping. If linearity is required, then setting $b(y) = vy$ for any fixed $v \in \mathbb{R}^2$ implies that $\mathscr{T}^*(y,u) + \mathscr{B}(y,u)$ is an adjoint operator in the sense of Definition 2.1.

It is not necessary in many applications to have a globally defined $\mathscr{T}^*$. The following theorem will lead to a sufficient condition for the existence of a locally defined adjoint operator.

Theorem 2.1 (Gray and Scherpen [22]). *Assume H is a Hilbert space and $U \subset H$ is any convex neighborhood of* 0. *Let $L: U \mapsto \mathbb{R}$ be a continuously Fréchet differentiable mapping on U with $L(0) = 0$. Then L has a factorization of the form*

$$L(u) = \langle a(u), u\rangle_H,$$

where $a: U \mapsto H$ is continuous on U, and for each $u \in U$ the dual mapping (*from the Riesz representation*) *is*

$$\begin{aligned}(a(u))^*: H &\mapsto \mathbb{R}\\ :\xi &\mapsto \int_0^1 (DL(tu))(\xi)\,\mathrm{d}t.\end{aligned}$$

This theorem can be viewed as a kind of infinite-dimensional version of the Fundamental Theorem of Integral Calculus. Its application in the nonlinear Hilbert adjoint existence theorem is as follows.

Theorem 2.2 (Gray and Scherpen [22]). *Suppose H_1 and H_2 are two Hilbert spaces and $U \subset H_1$ is any convex neighborhood of 0. Let $\mathscr{T}: U \mapsto H_2$ be a continuously Fréchet differentiable mapping on U such that $\mathscr{T}(0) = 0$. Then there exists a continuous mapping $\mathscr{T}^*: H_2 \times U \mapsto H_1$ with*

$$\langle \mathscr{T}(u), y\rangle_{H_2} = \langle u, \mathscr{T}^*(y,u)\rangle_{H_1}, \quad \forall u \in U, \quad \forall y \in H_2.$$

Specifically, $\mathscr{T}^*(y,u) = a_y(u)$ is such a mapping, where $a_y(\cdot)$ is defined for any fixed $y \in H_2$ by Theorem 2.1 with $L_y(u) = \langle \mathscr{T}(u), y\rangle_{H_2}$.

In [8,27,29,38,40] the characterization of the adjoint operator given in Theorem 2.2 (or more correctly Theorem 3.1) is basically used as the definition of a unique adjoint for a homogeneous operator. Eq. (2) is simply viewed as a property of this adjoint operator. In [38] the definition is further extended to handle homogeneous operators that depend on a single parameter $\varepsilon \in [0,1]$. In [8] nonhomogeneous operators with boundary conditions are also considered. In our case, the adjoint operator in Theorem 2.2 is just one of many possible solutions to Definition 2.1. Different definitions of adjoint operators can be found in [2,9]. In [2] a pseudo-adjoint operator is considered in the context of Lipschitz operators. The definition in [9] is introduced specifically for solving nonlinear partial differential equations by using a nonlinear semigroup generated by an accretive operator.

Eigen-structure of the nonlinear Hankel operator: In the theory of continuous-time linear systems, the system Hankel operator plays an important role in a number of realization problems. Interpretations both in terms of input–output mappings and state-space settings are available and have shown to be extremely useful in a number of applications, such as model reduction and system identification.

In the case of the nonlinear Hankel operator, primarily state-space notions have provided the useful tools [13–16,36]. We introduce here some background material from [13,14] which is later applied to spectral analysis problems in Section 4 in order to obtain a more complete theory in the input–output framework and to better relate it to the state-space setting. Consider a smooth time-invariant input-affine nonlinear control systems with no direct feed-through, i.e.,

$$\Sigma: \quad \begin{aligned} \dot{x} &= f(x) + g(x)u, \\ y &= h(x), \end{aligned} \tag{4}$$

where $u = (u_1, \ldots, u_m) \in \mathbb{R}^m$, $y = (y_1, \ldots, y_p) \in \mathbb{R}^p$, and $x = (x_1, \ldots, x_n)$ are local coordinates for a smooth state-space manifold denoted by M. Throughout it is assumed that the system has an isolated equilibrium. Without loss of generality, this equilibrium is taken to be at 0, i.e., $f(0) = 0$. It is also assumed that $h(0) = 0$. It is necessary that the system be well defined on the time interval $(-\infty, \infty)$. Finally, it is assumed throughout that

(A1) Σ is L_2-stable in the sense that $u \in L_2^m(-\infty, 0]$ implies that $\Sigma(u)$ restricted to $[0,\infty)$ is in $L_2^p[0,\infty)$.

The original definitions of the observability and controllability operators for Σ are given in [19,22] in terms of Chen–Fliess functional expansions [11]. But one can also employ state-space systems to describe them, specifically:

$$y = \mathcal{O}_\Sigma(x^0): \quad \begin{cases} \dot{x} = f(x) & x(0) = x^0, \\ y = h(x), \end{cases} \tag{5}$$

$$x^1 = \mathcal{C}_\Sigma(u): \quad \begin{cases} \dot{x} = f(x) + g(x)\mathcal{F}_-(u) & x(-\infty) = 0, \\ x^1 = x(0), \end{cases} \tag{6}$$

where $\mathcal{F}_- : L_2^m(-\infty,\infty) \to L_2^m(-\infty,0]$ denotes the time-flipping operator defined by

$$\mathcal{F}_-(u)(t) := \begin{cases} u(-t), & t \in (-\infty, 0], \\ 0, & t \in [0,\infty). \end{cases}$$

The Hankel operator $\mathcal{H}_\Sigma : L_2^m[0,\infty) \to L_2^p[0,\infty)$ for Σ is given by $\mathcal{H}_\Sigma := \Sigma\mathcal{F}_-$, and the identity $\mathcal{H}_\Sigma = \mathcal{O}_\Sigma\mathcal{C}_\Sigma$ was also proven in [19,22]. State-space descriptions of the corresponding adjoint operators can be found in [15,16]. In [13,14] the adjoint of the variational version of the Hankel operator has been shown to be useful for an eigen-structure analysis of the Hankel operator. These results are summarized next. In order to describe an eigen-structure of the Hankel operator, a state-space realization and corresponding pair of energy functions are employed as described below.

Definition 2.2. The observability function $L_o(x)$ and the controllability function $L_c(x)$ of Σ in (4) are defined by

$$L_o(x^0) := \frac{1}{2}\int_0^\infty \|y(t)\|^2 \, dt, \quad x(0) = x^0, \quad u(t) \equiv 0, \tag{7}$$

$$L_c(x^1) := \min_{\substack{u \in L_2^m[0,\infty) \\ x(-\infty)=0, x(0)=x^1}} \frac{1}{2}\int_{-\infty}^0 \|u(t)\|^2 \, dt. \tag{8}$$

It is assumed throughout that

(A2) There exist well-defined smooth observability and controllability functions L_o and L_c.

These functions are closely related to the observability and controllability operators above. In [33] they have been used for the definition of balanced realizations and singular value functions for nonlinear systems. They also fulfill corresponding Hamilton–Jacobi equations, in a similar way as the observability Gramian and the inverse of the controllability Gramian are solutions of a Lyapunov/Riccati equation. Assuming that Σ is Fréchet differentiable, and that $D\Sigma$ is L_2 input–output stable, then the following lemma was proven.

Lemma 2.1 (Fujimoto and Scherpen [13,14]). *If there exist $\lambda \in \mathbb{R}$ and a nonzero $x^0 \in \mathbb{R}^n$ such that*

$$\frac{\partial L_{\mathrm{o}}}{\partial x}(x^0) = \lambda \frac{\partial L_{\mathrm{c}}}{\partial x}(x^0), \tag{9}$$

then λ is the eigenvalue of the mapping $u \mapsto (D\mathscr{H}_\Sigma(u))^* \mathscr{H}_\Sigma(u)$ with corresponding eigen vector

$$v = \mathscr{C}_\Sigma^\dagger(x^0), \tag{10}$$

where $\mathscr{C}_\Sigma^\dagger : \mathbb{R}^n \to L_2^m[0,\infty)$ denotes the pseudo-inverse of $\mathscr{C}_\Sigma$ defined by

$$\mathscr{C}_\Sigma^\dagger(x^1) := \arg \min_{\mathscr{C}_\Sigma(u)=x^1} \|u\|_2. \tag{11}$$

This lemma relates the gradient of the controllability and observability functions to the eigenvalues of $(D\mathscr{H}_\Sigma(u))^* \mathscr{H}_\Sigma(u)$. The next result gives a more general parameterized eigen-structure of $(D\mathscr{H}_\Sigma(u))^* \mathscr{H}_\Sigma(u)$ in terms of energy level sets in the state-space and relates it to $\mathscr{H}_\Sigma^*(\mathscr{H}_\Sigma(u), u)$.

Theorem 2.3 (Fujimoto and Scherpen [13,14]). *Suppose the energy functions $L_{\mathrm{o}}(x)$ and $L_{\mathrm{c}}(x)$ are sufficiently smooth and that the Jacobian linearization of the system Σ has n distinct Hankel singular values. Then there exists locally $2n$ smooth singular value functions $\rho_i^j(c)$, $i \in \{1,2,\ldots,n\}$, $j \in \{+,-\}$ such that* $\min\{\rho_i^+(c), \rho_i^-(c)\} > \max\{\rho_{i+1}^+(c), \rho_{i+1}^-(c)\}$, $(\rho_{\max}(c) := \max\{\rho_1^+(c), \rho_1^-(c)\}$, $\rho_{\min}(c) := \min\{\rho_n^+(c),\ \rho_n^-(c)\})$, *and there exists parameterized vectors $x_i^j(c)$ satisfying*

$$L_{\mathrm{c}}(x_i^j(c)) = \frac{c^2}{2}, \quad L_{\mathrm{o}}(x_i^j(c)) = \frac{c^2(\rho_i^j(c))^2}{2}, \tag{12}$$

$$\frac{\partial L_{\mathrm{o}}}{\partial x}(x_i^j(c)) = \lambda_i^j(c) \frac{\partial L_{\mathrm{c}}}{\partial x}(x_i^j(c)) \tag{13}$$

with $\lambda_i^j(c) := (\rho_i^j(c))^2 + (\mathrm{d}(\rho_i^j(c))^2/\mathrm{d}c)c/2$. *Furthermore, when $u_i^j(c) := \mathscr{C}_\Sigma^\dagger(x_i^j(c))$, it follows that*

$$\langle u_i^j(c), \mathscr{H}_\Sigma^*(\mathscr{H}_\Sigma(u_i^j(c)), u_i^j(c))\rangle_{L_2^m} = (\rho_i^j(c))^2 \langle u_i^j(c), u_i^j(c)\rangle_{L_2^m} \tag{14}$$

and thus, the Hankel norm of the system is given by

$$\|\Sigma\|_H := \|\mathscr{H}_\Sigma\|_{L_2^m} = \sup_{u \in L_2^m(-\infty,0]} \frac{\|\mathscr{H}_\Sigma(u)\|_2}{\|u\|_2} = \sup_{c>0}\{\max\{\rho_1^+(c), \rho_1^-(c)\}\}. \tag{15}$$

Remark 2.1. From the above theorem the following local linear interpretations are possible:

(1) $\rho_i^j(0)$ is a Hankel singular value of the linearized system with $\rho_i^-(0) = \rho_i^+(0)$.

(2) $\lambda_i^j(0)$ is a squared Hankel singular value of the linearized system with $\lambda_i^-(0) = \lambda_i^+(0)$.

(3) $x_i^j(0)$ is an eigen vector of PQ, where P and Q are the controllability and observability Gramians, respectively, of the linearized system with $x_i^-(0) = x_i^+(0)$.
(4) $u_i^j(0)$ is an eigen vector of $\mathscr{H}_\Sigma^* \mathscr{H}_\Sigma$ for the linearized system with $u_i^-(0) = u_i^+(0)$.

With the necessary background material in place, we now proceed to the first main topic of the paper.

3. Adjoint properties and examples

While a useful device in many circumstances, a nonlinear Hilbert adjoint operator does not share all of the familiar properties associated with linear adjoint operators. Consider any normed set of linear operators B defined on $L_2[0,\infty)$ as a Banach algebra with composition product $(\mathscr{S}, \mathscr{T}) \mapsto \mathscr{S}\mathscr{T}$. B is said to constitute a C^*-algebra if it is equipped with an *adjoint map* (or *involution*) $\mathscr{T} \mapsto \mathscr{T}^*$ such that for all $\mathscr{S}, \mathscr{T} \in B$ and any $\alpha \in \mathbb{R}$, the following properties are satisfied [39]:

(i) (linearity) $(\alpha\mathscr{S} + \mathscr{T})^* = \alpha\mathscr{S}^* + \mathscr{T}^*$;
(ii) (product-reversal) $(\mathscr{S}\mathscr{T})^* = \mathscr{T}^*\mathscr{S}^*$;
(iii) (double adjoint) $(\mathscr{T}^*)^* = \mathscr{T}$; and
(iv) (C^*-identity) $\|\mathscr{T}\|^2 = \|\mathscr{T}^*\mathscr{T}\|$.

In this section the appropriate extensions of these fundamental properties are presented for the nonlinear Hilbert adjoint. Interspersed in the presentation is a collection of simple examples meant to illustrate the main ideas.

The linearity property (i) is an immediate result which follows from the bi-linearity of the inner product and the interpretation that equality here implies belonging to the same equivalence class. An interesting side issue is that when an adjoint operator exists, i.e., fulfills identity (2), it follows from Theorem 2.2 that there always exists an explicit form which is linear in the first argument.

Theorem 3.1. *Suppose H_1 and H_2 are two Hilbert spaces and $U \subset H_1$ is any convex neighborhood of 0. Let $\mathscr{T} : U \mapsto H_2$ be a continuously Fréchet differentiable mapping on U such that $\mathscr{T}(0) = 0$. Then the mapping*

$$\mathscr{T}^*(y,u) = \int_0^1 (D\mathscr{T}(tu))^*(y)\,\mathrm{d}t$$

is a suitable Hilbert adjoint of $\mathscr{T}$ on $H_2 \times U$.

Proof. For any $y \in H_2$, define the scalar-valued mapping on U:

$$L_y(u) = \langle \mathscr{T}(u), y\rangle_{H_2} \equiv \langle u, \mathscr{T}^*(y,u)\rangle_{H_1}.$$

Next, observe that for any fixed $u \in U$ and $t \in [0,1]$ it follows that

$$\begin{aligned} DL_y(tu)(\xi) &= \langle D\mathscr{T}(tu)(\xi), y\rangle_{H_2} \\ &= \langle \xi, (D\mathscr{T}(tu))^*(y)\rangle_{H_1}, \quad \forall \xi \in H_1. \end{aligned}$$

Thus,

$$L_y(u) = \int_0^1 (DL_y(tu))(u)\,\mathrm{d}t$$
$$= \int_0^1 \langle u, (D\mathscr{T}(tu))^*(y)\rangle_{H_1}\,\mathrm{d}t$$
$$= \left\langle u, \int_0^1 (D\mathscr{T}(tu))^*(y)\,\mathrm{d}t \right\rangle_{H_1}$$

and the conclusion follows directly. □

Observe that in this form above, $\mathscr{T}^*(y,u)$ is linear in y since $(D\mathscr{T}(tu))^*(\cdot)$ is the adjoint of a linear operator, i.e., the familiar Hilbert adjoint. Thus, it is also immediate that $\mathscr{T}^*(0,u)=0$, $\forall u \in U$. Unfortunately, Example 2.1 demonstrates that this linearity property is still not sufficient to provide any uniqueness features.

Example 3.1. For any finite $T>0$ and positive integer m, the Banach space $L_4^m[0,T]$ can be viewed as a convex open subset of $L_2^m[0,T]$ containing the zero function. With $U = L_4^m[0,T]$, the mapping

$$\mathscr{T} : U \mapsto L_2[0,T],$$
$$: u \mapsto u^{\mathrm{T}} u$$

is then well defined, continuously Fréchet differentiable, and satisfies the identity $\mathscr{T}(0)=0$. One form of the adjoint operator can be immediately extracted using the definition. Specifically,

$$\langle \mathscr{T}(u), y\rangle_{L_2} = \int_0^T u^{\mathrm{T}}(\tau)u(\tau)y(\tau)\,\mathrm{d}\tau,$$
$$= \langle u, uy\rangle_{L_2^m}$$

and thus, a suitable adjoint is given by $\mathscr{T}^*(y,u)=uy$. This same adjoint form can also be computed using Theorem 3.1:

$$D\mathscr{T}(u) = 2u^{\mathrm{T}},$$
$$(D\mathscr{T}(tu))^*(y) = 2tuy,$$
$$\mathscr{T}^*(y,u) = \int_0^1 (D\mathscr{T}(tu))^*(y)\,\mathrm{d}t = uy.$$

Example 3.2. Consider a Hammerstein integral operator defined on a set $U \subset L_2^m[0,\infty)$:

$$\mathscr{S} : U \mapsto L_2^p[0,\infty)$$
$$: u \mapsto \int_0^\infty K(\tau,s)f(u(s))\,\mathrm{d}s,$$

where K is a suitable continuous kernel function, and each component function of f is C^1 with $f(0)=0$. Then applying Theorem 3.1 it follows that

$$(D\mathscr{S}(u))(\xi) = \int_0^\infty K(\tau,s)\frac{\mathrm{d}f}{\mathrm{d}u}(u(s))\xi(s)\,\mathrm{d}s,$$

$$(D\mathscr{S}(u))^*(y) = \int_0^\infty \frac{\mathrm{d}f}{\mathrm{d}u}^{\mathrm{T}}(u(s))K^{\mathrm{T}}(\tau,s)y(\tau)\,\mathrm{d}\tau,$$

$$\mathscr{S}^*(y,u) = \int_0^\infty \underbrace{\left[\int_0^1 \frac{\mathrm{d}f}{\mathrm{d}u}(tu(s))\,\mathrm{d}t\right]^{\mathrm{T}}}_{F(u(s))} K^{\mathrm{T}}(\tau,s)y(\tau)\,\mathrm{d}\tau$$

$$= F^{\mathrm{T}}(u)S_{\mathrm{L}}^*(y),$$

where the matrix-valued function $F(\cdot)$ satisfies the identity $f(x)=F(x)x$ on a convex neighborhood of 0, and S_{L}^* denotes the usual adjoint operator for the linear integral operator with kernel K. Some specific examples include the familiar linear time-invariant case

$$\mathscr{S}_{\mathrm{LTI}}(u) = \int_0^\infty \mathrm{e}^{A(\tau-s)}u(s)\,\mathrm{d}s,$$

where A is Hurwitz and $f(x)=x$, and thus a suitable adjoint is given by

$$\mathscr{S}_{\mathrm{LTI}}^*(y,u) = \int_0^\infty \mathrm{e}^{A^{\mathrm{T}}(\tau-s)}y(\tau)\,\mathrm{d}\tau.$$

Also consider the SISO Frequency Modulated (FM) system

$$\mathscr{S}_{\mathrm{FM}}(u) = \frac{1}{\pi}\int_0^\infty \mathrm{e}^{A(\tau-s)}\sin(\pi u(s))\,\mathrm{d}s,$$

where a suitable adjoint is given by

$$\mathscr{S}_{\mathrm{FM}}^*(y,u) = \mathrm{sinc}(u)\int_0^\infty \mathrm{e}^{A^{\mathrm{T}}(\tau-s)}y(\tau)\,\mathrm{d}\tau$$

$$= \mathrm{sinc}(u)S_{\mathrm{LTI}}^*(y).$$

In order to address the product-reversal property (ii), one must first define the sense in which operators can be composed when adjoint operators are present. The situation is more complicated than the familiar case since the domain of an adjoint operator is not simply the codomain of the original operator. For example, consider the Hilbert spaces H_i, $i=1,2,3$, the operators

$$\begin{aligned} \mathscr{T} &: H_1 \mapsto H_2 & \mathscr{S} &: H_2 \mapsto H_3 \\ &: u \mapsto w & &: w \mapsto y \end{aligned}$$

and the corresponding adjoints

$$\begin{aligned} \mathscr{T}^* &: H_2 \times H_1 \mapsto H_1 & \mathscr{S}^* &: H_3 \times H_2 \mapsto H_2 \\ &: (w,u) \mapsto \bar{u} & &: (y,w) \mapsto \bar{w}. \end{aligned}$$

Clearly the composition and its adjoint

$$\begin{aligned} \mathscr{S}\mathscr{T} &: H_1 \mapsto H_3 \qquad & (\mathscr{S}\mathscr{T})^* &: H_3 \times H_1 \mapsto H_1 \\ &: u \mapsto y & &: (y,u) \mapsto \bar{u}. \end{aligned}$$

are well defined, but no *direct* composition like $\mathscr{T}^*\mathscr{T}$ or $\mathscr{T}^*\mathscr{S}^*$ is possible as in the classic setting. Still some *formal* compositions can be defined which have great utility in a variety of situations.

Definition 3.1. Let H_i, $i = 1,2,3$, be a collection of Hilbert spaces. Assume $\mathscr{T} : H_1 \mapsto H_2$ and $\mathscr{S} : H_2 \mapsto H_3$ are two operators with well-defined adjoint operators. Define the following operator products:

$$\begin{aligned} (\mathscr{S}^*\mathscr{T})_1 &: H_1 \times H_2 \mapsto H_2 \quad [\text{when } H_2 = H_3] \\ &: (u,w) \mapsto \mathscr{S}^*(\mathscr{T}(u),w) \\ (\mathscr{S}^*\mathscr{T})_2 &: H_3 \times H_1 \mapsto H_1 \\ &: (y,u) \mapsto \mathscr{S}^*(y,\mathscr{T}(u)). \end{aligned}$$

The main application of this definition in this paper is in regards to the product-reversal property.

Theorem 3.2 (Product-reversal). *Let H_i, $i = 1,2,3$, be a collection of Hilbert spaces. Assume $\mathscr{T} : H_1 \mapsto H_2$ and $\mathscr{S} : H_2 \mapsto H_3$ are two operators with well-defined adjoint operators. Then the following identity holds*:

$$(\mathscr{S}\mathscr{T})^* \cong (\mathscr{T}^*(\mathscr{S}^*\mathscr{T})_2)_1.$$

Proof. The claim follows straightforwardly from the defining property (2). Observe that for any $(y,u) \in H_3 \times H_1$:

$$\begin{aligned} \langle u, (\mathscr{S}\mathscr{T})^*(y,u)\rangle_{H_1} &= \langle \mathscr{S}\mathscr{T}(u), y\rangle_{H_3} \\ &= \langle \mathscr{T}(u), \mathscr{S}^*(y,\mathscr{T}(u))\rangle_{H_2} \\ &= \langle u, \mathscr{T}^*(\mathscr{S}^*(y,\mathscr{T}(u)),u)\rangle_{H_1} \\ &= \langle u, (\mathscr{T}^*(\mathscr{S}^*\mathscr{T})_2)_1(y,u)\rangle_{H_1}. \qquad \square \end{aligned}$$

In order to compute adjoints of general adjoint operators for the double adjoint property (iii), the concept of a partial adjoint operator is needed. The idea is based on a direct generalization of identity (2).

Definition 3.2. For any set of Hilbert spaces H_i, $i = 1,\ldots,m+1$ and an operator

$$\begin{aligned} \mathscr{U} &: H_1 \times H_2 \times \cdots \times H_m \mapsto H_{m+1} \\ &: (u_1,u_2,\ldots,u_m) \mapsto y, \end{aligned} \tag{16}$$

a jth partial adjoint of $\mathscr{U}$ is any mapping of the form

$$\mathscr{U}^{*_j} : H_{m+1} \times H_1 \times H_2 \times \cdots \times H_m \mapsto H_j,$$

where

$$\langle \mathscr{U}(u_1, u_2, \ldots, u_m), y \rangle_{H_{m+1}} = \langle u_j, \mathscr{U}^{*_j}(y, u_1, u_2, \ldots, u_m) \rangle_{H_j}, \quad \forall u_i \in H_i, \quad y \in H_{m+1}$$

for $i = 1, \ldots, m$.

These definitions produce the following double adjoint identities.

Theorem 3.3 (Double adjoints). *Let H_1 and H_2 be two Hilbert spaces and $\mathscr{T} : H_1 \mapsto H_2$ be an operator with a well defined adjoint. Then it follows that*

$$(\mathscr{T}^*)^{*_1}(\bar{u}, y, u)|_{\bar{u}=u} \cong \mathscr{T}(u),$$

$$(\mathscr{T}^*)^{*_2}(\bar{u}, y, u)|_{\bar{u}=u} \cong \mathscr{T}^*(y, u)$$

for all $u \in H_1, y \in H_2$, assuming all the partial adjoints exist.

Proof. With respect to the first identity, observe that the first partial adjoint of $\mathscr{T}^*(y, u)$ fulfills

$$\begin{aligned}\langle y, (\mathscr{T}^*)^{*_1}(\bar{u}, y, u) \rangle|_{\bar{u}=u} &= \langle \mathscr{T}^*(y, u), \bar{u} \rangle|_{\bar{u}=u} \\ &= \langle y, \mathscr{T}(u) \rangle.\end{aligned}$$

For the second partial adjoint of $\mathscr{T}^*(y, u)$,

$$\begin{aligned}\langle u, (\mathscr{T}^*)^{*_2}(\bar{u}, y, u) \rangle|_{\bar{u}=u} &= \langle \mathscr{T}^*(y, u), \bar{u} \rangle|_{\bar{u}=u} \\ &= \langle u, \mathscr{T}^*(y, u) \rangle. \qquad \square\end{aligned}$$

One application of this theorem is in regards to testing for self-adjointness.

Definition 3.3. Let H be a Hilbert space and $\mathscr{S} : H \mapsto H$ be a mapping with a well defined adjoint operator $\mathscr{S}^* : H \times H \mapsto H$. $\mathscr{S}$ is self-adjoint if

$$\mathscr{S}^*(\bar{u}, u)|_{\bar{u}=u} \cong \mathscr{S}(u), \quad \forall u \in H.$$

Observe that an operator like $\mathscr{T}^*\mathscr{T}(u) := (\mathscr{T}^*\mathscr{T})_1(\bar{u}, u)|_{\bar{u}=u} = \mathscr{T}^*(\mathscr{T}(u), u)$ is always self-adjoint since one may write in terms of the first partial adjoint

$$\begin{aligned}\langle \mathscr{T}^*(\mathscr{T}(u), u), \bar{u} \rangle_H &= \langle \mathscr{T}(u), (\mathscr{T}^*)^{*_1}(\bar{u}, \mathscr{T}(u), u) \rangle_H \\ &= \langle u, \mathscr{T}^*((\mathscr{T}^*)^{*_1}(\bar{u}, \mathscr{T}(u), u), u) \rangle_H,\end{aligned}$$

or in terms of the second partial adjoint

$$\langle \mathscr{T}^*(\mathscr{T}(u), u), \bar{u} \rangle_H = \langle u, (\mathscr{T}^*)^{*_2}(u, \mathscr{T}(u), \bar{u}) \rangle_H.$$

By definition it then follows that

$$(\mathscr{T}^*\mathscr{T})^*(\bar{u}, u) \cong \mathscr{T}^*((\mathscr{T}^*)^{*_1}(\bar{u}, \mathscr{T}(u), u), u),$$

$$(\mathscr{T}^*\mathscr{T})^*(\bar{u}, u) \cong (\mathscr{T}^*)^{*_2}(u, \mathscr{T}(u), \bar{u}).$$

In either case, the identities in Theorem 3.3 yield the required property

$$(\mathscr{T}^*\mathscr{T})^*(\bar{u},u)|_{\bar{u}=u} \cong (\mathscr{T}^*\mathscr{T})(u).$$

Example 3.3. Consider the operator and its suitable choice of nonlinear Hilbert adjoint given in Example 3.1, where now $m=1$. Then

$$\mathscr{T}^*(y,u)|_{y=u} = uy|_{y=u} = u^2 = \mathscr{T}(u).$$

So $\mathscr{T}$ is self-adjoint.

Example 3.4. Reconsider Example 3.2 where $m=p=1$. Even in this SISO case, the corresponding Hammerstein operator is rarely self-adjoint since:

$$\begin{aligned}\mathscr{S}^*(y,u)|_{y=u} &= F(u)S_{\mathrm{L}}^*(u)\\ &\neq \mathscr{S}(u). \qquad \square\end{aligned}$$

The final property under consideration is the "C^*-identity" (iv). Unlike the linear case, at present only an inequality is known to relate the two norms in question.

Theorem 3.4 (C^*-inequality). *Let H_1 and H_2 be Hilbert spaces. Assume $\mathscr{T}: H_1 \mapsto H_2$ is a bounded operator with a well-defined adjoint operator. Then the following inequality holds*:

$$\|\mathscr{T}\|^2 \leqslant \|\mathscr{T}^*\mathscr{T}\|.$$

Proof. For any fixed $u \in H_1$ and employing the Schwarz inequality,

$$\begin{aligned}\|\mathscr{T}(u)\|_{H_2}^2 &= \langle \mathscr{T}(u), \mathscr{T}(u)\rangle_{H_2}\\ &= \langle \mathscr{T}^*(\mathscr{T}(u),u),u\rangle_{H_1}\\ &= \langle \mathscr{T}^*\mathscr{T}(u),u\rangle_{H_1}\\ &\leqslant \|\mathscr{T}^*\mathscr{T}(u)\|_{H_1}\|u\|_{H_1}\end{aligned}$$

Dividing both sides by $\|u\|_{H_1}^2$ and taking the supremum over all $u \neq 0$ gives the final result. $\square$

The above theorems show that one almost has a complete nonlinear extension of a linear adjoint map defined on a C^* algebra, except for the equality in property (iv).

The section is concluded by considering how the Fréchet derivative interacts with nonlinear Hilbert adjoints. This is important because of its relationship to the eigen-structure of the Hankel operator described in Section 2 and its application to the spectral analysis in Section 4. Given an operator $\mathscr{U}$ of the form (16), its Fréchet derivative with respect to u_j at $(u_1,u_2,\ldots,u_m)$ is denoted by $D_j\mathscr{U}(u_1,u_2,\ldots,u_m)$. The situation is greatly simplified by the fact that $D_j\mathscr{U}(u_1,u_2,\ldots,u_m)$ is a linear operator defined on H_j.

Theorem 3.5. *Let H_1 and H_2 be two Hilbert spaces and $\mathscr{T}: H_1 \mapsto H_2$ be an operator with a well defined Hilbert adjoint. Assuming both $\mathscr{T}$ and $\mathscr{T}^*$ are Fréchet differentiable, then the following identities hold*:

(1) $(D_1\mathscr{T}^*(y,u))^*(u) = \mathscr{T}(u)$,

(2) $(D_2\mathscr{T}^*(y,u))^*(u) = (D\mathscr{T}(u))^*(y) - \mathscr{T}^*(y,u)$,

(3) $(D(\mathscr{T}^*\mathscr{T}(u)))^*(u) = 2(D\mathscr{T}(u))^*(\mathscr{T}(u)) - \mathscr{T}^*\mathscr{T}(u)$.

Proof. (1) For any $u \in H_1$ and $\xi, y \in H_2$ observe that

$$D_y\langle \mathscr{T}^*(y,u), u\rangle_{H_1}(\xi) = D_y\langle y, \mathscr{T}(u)\rangle_{H_2}(\xi),$$

$$\langle D_1\mathscr{T}^*(y,u)(\xi), u\rangle_{H_1} = \langle \xi, \mathscr{T}(u)\rangle_{H_2},$$

$$\langle \xi, (D_1\mathscr{T}^*(y,u))^*(u)\rangle_{H_2} = \langle \xi, \mathscr{T}(u)\rangle_{H_2}.$$

(2) Similarly, for any $u, \xi \in H_1$ and $y \in H_2$

$$D_u\langle \mathscr{T}^*(y,u), u\rangle_{H_1}(\xi) = D_u\langle y, \mathscr{T}(u)\rangle_{H_2}(\xi),$$

$$\langle D_2\mathscr{T}^*(y,u)(\xi), u\rangle_{H_1} + \langle \mathscr{T}^*(y,u), \xi\rangle_{H_1} = \langle y, D\mathscr{T}(u)(\xi)\rangle_{H_2},$$

$$\langle \xi, (D_2\mathscr{T}^*(y,u))^*(u)\rangle_{H_1} = \langle \xi, (D\mathscr{T}(u))^*(y)\rangle_{H_2} - \langle \xi, \mathscr{T}^*(y,u)\rangle_{H_1}.$$

(3) First observe that for any $u, \xi \in H_1$

$$\langle u, D(\mathscr{T}^*(\mathscr{T}(u),u))(\xi)\rangle_{H_1} = \langle u, D_1(\mathscr{T}^*(\mathscr{T}(u),u))(D\mathscr{T}(u)(\xi)) + D_2(\mathscr{T}^*(\mathscr{T}(u),u))(\xi)\rangle_{H_1}.$$

Now, using the previous two identities it follows that

$$\begin{aligned}
&(D(\mathscr{T}^*\mathscr{T}(u)))^*(u)\\
&\quad = (D_1(\mathscr{T}^*(\mathscr{T}(u),u))(D\mathscr{T}(u)))^*(u) + (D_2(\mathscr{T}^*(\mathscr{T}(u),u)))^*(u)\\
&\quad = (D\mathscr{T}(u))^*(D_1(\mathscr{T}^*(\mathscr{T}(u),u)))^*(u) + (D_2(\mathscr{T}^*(\mathscr{T}(u),u)))^*(u)\\
&\quad = (D\mathscr{T}(u))^*(\mathscr{T}(u)) - \mathscr{T}^*(\mathscr{T}(u),u) + (D\mathscr{T}(u))^*(\mathscr{T}(u)). \qquad \square
\end{aligned}$$

4. Towards spectral analysis of $\mathscr{H}_\Sigma^*\mathscr{H}_\Sigma$

Spectral theory for nonlinear operators is a diverse subject with substantial roots going back to at least the late 1960's [6]. The proliferation of definitions and approaches (see, for example, [3,12,17,18,23,24,31] is partly due to the fact that no single definition completely characterizes the original operator as in the linear case. In this section, we outline an additional approach to defining a nonlinear spectrum motivated by the nature of our application and the notion of the C^1-spectrum essentially introduced in [30].

Definition 4.1. Let E be a Banach space and $\mathscr{S}: E \to E$ be an operator that is continuously Fréchet differentiable on E. The C^1-spectrum of $\mathscr{S}$, $\sigma^1(\mathscr{S})$, is the set of all complex numbers λ such that $\mathscr{S} - \lambda I$ is not a diffeomorphism on E.

For a linear operator $\mathscr{S}$, this definition reduces to the usual definition of a spectrum. The following result from the analysis of the C^1-spectrum in [3] is relevant to our study:

Theorem 4.1 (Appell and Dörfner [3]). *Let $\mathscr{S}$ be an operator as described in Definition* 4.1, *then*

$$\sigma^1(\mathscr{S}) = \pi(\mathscr{S}) \cup \bigcup_{u \in E} \sigma(D\mathscr{S}(u)),$$

where $\pi(\mathscr{S})$ denotes the set of all λ such that $\mathscr{S} - \lambda I$ is not proper (*in the sense of* [3]), *and $\sigma(\mathscr{A})$ denotes the usual spectrum of a bounded linear operator $\mathscr{A}$.*

This theorem reveals that the C^1-spectrum of a nonlinear operator directly involves the Fréchet derivative of the operator, i.e., $\sigma(D\mathscr{S}(u))$ is an important part of the C^1-spectrum. Since our problem is to extend the singular value definitions into the nonlinear setting, and in particular for a Hankel operator, it is the spectrum of the Fréchet derivative of the operator $\mathscr{H}_\Sigma^* \mathscr{H}_\Sigma(u)$ that is relevant. The following corollary of Theorems 2.3 and 3.5 reveals some information about $\sigma(D(\mathscr{H}_\Sigma^* \mathscr{H}_\Sigma))$.

Corollary 4.1. *In the context of Theorem* 2.3, *the following relation holds*:

$$\langle (D(\mathscr{H}_\Sigma^* \mathscr{H}_\Sigma)(u_i^j(c)))(u_i^j(c)), u_i^j(c) \rangle = (2\lambda_i^j(c) - (\rho_i^j(c))^2) \langle u_i^j(c), u_i^j(c) \rangle.$$

Proof. Applying Theorem 3.5, property 3, gives directly

$$\begin{aligned} &\langle (D(\mathscr{H}_\Sigma^* \mathscr{H}_\Sigma)(u_i^j(c)))(u_i^j(c)), u_i^j(c) \rangle \\ &\quad = 2\langle (D\mathscr{H}_\Sigma(u_i^j(c)))(\mathscr{H}_\Sigma(u_i^j(c))), u_i^j(c) \rangle - \langle \mathscr{H}_\Sigma^* \mathscr{H}_\Sigma(u_i^j(c)), u_i^j(c) \rangle. \end{aligned}$$

Then using Theorem 2.3, the result immediately follows. □

Note that the above corollary does not directly characterize $\sigma(D(\mathscr{H}_\Sigma^* \mathscr{H}_\Sigma))$, but it yields an eigen-equation within an inner product identity. In the general nonlinear setting, it is not possible to (exactly) extract the eigen-equation from this identity, i.e., every number in $\sigma(D(\mathscr{H}_\Sigma^* \mathscr{H}_\Sigma))$ fulfills the above equation, but numbers not in $\sigma(D(\mathscr{H}_\Sigma^* \mathscr{H}_\Sigma))$ can satisfy it too. In [3], other types of spectra are discussed, mainly for operators that are not C^1. The operators of the form $\mathscr{T}^*(\mathscr{T}(u), u)$ considered here are generally C^1. Furthermore, our original interest in the singular value structure of the Hankel operator is related to the state-space notions of the controllability and observability energy functions and the inner product relations with their respective operators (see [19,22,35]). These facts, together with the observation that our nonlinear Hilbert adjoint is defined within an inner product, motivates one to include the inner product structure directly into a spectrum definition.

Definition 4.2. Let E be a Hilbert space and consider an operator $\mathscr{S}: E \to E$. Then the inner product spectrum is defined as

$$\sigma_{\mathrm{i}p}(\mathscr{S}) = \{\lambda\colon \exists p \neq 0 \text{ with } \langle (\mathscr{S} - \lambda I)(p), p\rangle_E = 0\}.$$

It follows immediately when $\mathscr{S}$ is continuously Fréchet differentiable on E that $\sigma_{\mathrm{i}p}(\mathscr{S}) \subset \sigma^1(\mathscr{S})$. Furthermore, in the case of a linear operator $\mathscr{S}(p) = Ap$ with $A^{\mathrm{T}} = A$ and $E = \mathbb{R}^n$, it is easily verified that $\sigma_{\mathrm{i}p}(\mathscr{S}) = \mathrm{Range}(\mathscr{R}_{\mathscr{S}}(p))$, where $\mathscr{R}_{\mathscr{S}}$ is the Rayleigh quotient of $\mathscr{S}$ defined as

$$\mathscr{R}_{\mathscr{S}}(p) = \frac{p^{\mathrm{T}}Ap}{p^{\mathrm{T}}p}.$$

It is known in this case that $\sigma_{\mathrm{i}p}(A) = [\lambda_{min}(A), \lambda_{max}(A)] \subset \mathbb{R}$, where λ_{min} (λ_{max}) denotes the smallest (largest) eigenvalue of A. The obvious extension of the Rayleigh quotient for nonlinear maps is then

$$\begin{aligned} \mathscr{R}_{\mathscr{S}} &: E \to \mathbb{R} \\ &: p \mapsto \frac{\langle p, \mathscr{S}(p)\rangle_E}{\langle p, p\rangle_E}, \end{aligned}$$

and it straightforwardly follows that $\sigma_{\mathrm{i}p}(\mathscr{S}) = \mathrm{Range}(\mathscr{R}_{\mathscr{S}})$. When $\mathscr{S}$ is homogeneous, the range of this Rayleigh quotient is a subset of the numerical range defined in [41]. Furthermore, it is related to the numerical range $W(\mathscr{S}, \mathscr{T})$ as defined in [7] for positively homogeneous operators $\mathscr{S}$ and $\mathscr{T}$ of degree k on the unit sphere $S_1(0)$ in E. Specifically, if $\mathscr{T} = I$ and $\mathscr{S}$ is positively homogeneous of degree k, then $\mathscr{R}_{\mathscr{S}}(S_1(0)) = W(\mathscr{S}, I)$.

In the case of a compact linear operator $\mathscr{A}: E \to E$, it is known that

$$\sigma_{\mathrm{i}p}(\mathscr{A}^*\mathscr{A}) = (0, \tau_1^2],$$

where τ_1 is the largest singular value of $\mathscr{A}$ [26]. If $\mathrm{rank}(\mathscr{A}^*\mathscr{A}) = n < \infty$, then this result can be further refined to

$$\sigma_{\mathrm{i}p}(\mathscr{A}^*\mathscr{A}) = [\tau_n^2, \tau_1^2],$$

where τ_n is the smallest nonzero singular value of $\mathscr{A}$. In the case of the nonlinear system Σ given by (4) with corresponding Hankel operator $\mathscr{H}_\Sigma$, it follows immediately that $(\rho_i^j(c))^2 \in \sigma_{\mathrm{i}p}(\mathscr{H}_\Sigma^*\mathscr{H}_\Sigma)$ for all $i \in \{1, 2, \ldots, n\}$ and $j \in \{+, -\}$. Furthermore,

$$\|\Sigma\|_H^2 = \sup_{c>0} \{\max\{\rho_1^{+^2}(c), \rho_1^{-^2}(c)\}\} = \sup\{\sigma_{\mathrm{i}p}(\mathscr{H}_\Sigma^*\mathscr{H}_\Sigma)\}$$

and

$$\inf_{c>0} \{\min\{\rho_n^{+2}(c), \rho_n^{-2}(c)\}\} = \inf\{\sigma_{\mathrm{i}p}(\mathscr{H}_\Sigma^*\mathscr{H}_\Sigma)\},$$

where $\|\Sigma\|_H$ is the Hankel norm of Σ.

The following example illustrates some of the Hankel operator theory as it appears in this paper.

Example 4.1. Consider the following state-space system:

$$\Sigma: \quad \begin{cases} \dot{z}_1 = -z_1 + z_1 z_2^2 + u_1\sqrt{2}, \\ \dot{z}_2 = -z_2 - z_2^3 + u_2\sqrt{2 - 2z_1^2 + 2z_2^2}, \\ y_1 = 2z_1, \\ y_2 = \sqrt{2}z_2, \end{cases}$$

where $z \in W = \{z | z_1^2 < 1\}$. Then it can be shown that (see [34]) the controllability and observability function are given by

$$\tilde{L}_c(z) = \frac{1}{2} z^{\mathrm{T}} z, \quad \tilde{L}_0(z) = \frac{1}{2} z^{\mathrm{T}} \begin{pmatrix} 2 & 0 \\ 0 & 1 + z_1^2 \end{pmatrix} z.$$

In order to obtain the $\lambda_i^j(\cdot)$'s in (13), one must compute the solutions of

$$\begin{cases} 2z_1 + z_1 z_2^2 = \lambda(c) z_1, \\ z_2 + z_1^2 z_2 = \lambda(c) z_2, \\ z_1^2 + z_2^2 = c^2. \end{cases}$$

Note that $z \in W$ implies that $c < 1$, so that $z_1^2 = \frac{1}{2}(1 + c^2)$ and $z_2^2 = \frac{1}{2}(c^2 - 1)$ have no real solution. Thus, there remains two possibilities:

$$x_1^+(c) = \begin{pmatrix} c \\ 0 \end{pmatrix}, \quad \lambda_1^+(c) = 2,$$

$$x_1^-(c) = \begin{pmatrix} -c \\ 0 \end{pmatrix}, \quad \lambda_1^-(c) = 2,$$

$$x_2^+(c) = \begin{pmatrix} 0 \\ c \end{pmatrix}, \quad \lambda_2^+(c) = 1,$$

$$x_2^-(c) = \begin{pmatrix} 0 \\ -c \end{pmatrix}, \quad \lambda_2^-(c) = 1.$$

In [15,16] it is shown that

$$c\lambda_i^j(c) = \frac{\mathrm{d}}{\mathrm{d}c} \left(\rho_i^{j^2}(c) \frac{c^2}{2} \right),$$

and thus it follows that

$$\rho_1^{\pm^2}(c) = 2, \quad \rho_2^{\pm^2}(c) = 1 \Rightarrow \sigma_{\mathrm{i}p}(\mathscr{H}_\Sigma^* \mathscr{H}_\Sigma) = [1, 2] \Rightarrow \|\Sigma\|_H = \sqrt{2}.$$

5. Conclusions and future research

The existing notion of a nonlinear Hilbert adjoint was further developed by exhibiting its basic properties as extensions of those for adjoints in a C^*-algebra. Then, after defining self-adjointness in this context, nonlinear Hankel singular value analysis was performed using a new type of spectrum that directly incorporates the inner product structure. In the future more connections between this theory and the existing nonlinear Hankel theory will be explored further, especially its application to state-space model reduction.

Acknowledgements

This research was supported in part by a travel grant from the Netherlands Organization for Scientific Research (NWO), file number R62-554.

References

[1] R. Abraham, J.E. Marsden, T. Ratiu, Manifolds, Tensor Analysis, and Applications, 2nd Edition, Springer, New York, 1988.
[2] J. Appell, Ein merkwürdiges sepktrum für nichtlineare operatoren, Math. Bohemica 124 (2–3) (1999) 221–229.
[3] J. Appell, M. Dörfner, Some spectral theory for nonlinear operators, Nonlinear Anal. Theory Methods Appl. 28 (1997) 1955–1976.
[4] J. Batt, Nonlinear compact mappings and their adjoints, Math. Ann. 189 (1970) 5–25.
[5] J.G.F. Belinfante, B. Kolman, A Survey of Lie Groups and Lie Algebras with Applications and Computational Methods, SIAM, Philadelphia, 1972.
[6] F.E. Browder, Nonlinear eigenvalue problems and Galerkin approximations, Bull. Amer. Math. Soc. 74 (1968) 651–656.
[7] V. Burýsková, Some properties of nonlinear adjoint operators, Rocky Mountain J. Math. 28 (1) (1998) 41–59.
[8] D.G. Cacuci, R.B. Perez, V. Protopopescu, Duals and propagators: a canonical formalism for nonlinear equations, J. Math. Phys. 29 (2) (1988) 353–361.
[9] V. Caselles, Duality and nonlinear equations governed by accretive operators, Anal. Non Linéaire Publ. Math. Fac. Sci. Besançon 12 (1990) 7–25.
[10] P.E. Crouch, A.J. van der Schaft, Variational and Hamiltonian Control Systems, Lecture Notes in Control Information Science, Vol. 101, Springer, Berlin, 1987.
[11] M. Fliess, M. Lamnabhi, F. Lamnabhi-Lagarrigue, An algebraic approach to nonlinear functional expansions, IEEE Trans. Circuits Syst. 30 (1983) 554–570.
[12] G. Fournier, M. Martelli, Eigenvectors for nonlinear maps, Topol. Methods Nonlinear Anal. 2 (1993) 203–224.
[13] K. Fujimoto, J.M.A. Scherpen, Eigen structure of nonlinear Hankel operators, in: A. Isidori, F. Lamnabhi-Laguarrigue, W. Respondek (Eds.), Nonlinear Control in the Year 2000, Lecture Notes in Information Science, Vol. 1, Springer, Berlin, 2000, pp. 385–398.
[14] K. Fujimoto, J.M.A. Scherpen, Nonlinear input-normal realizations based on the differential eigenstructure of Hankel operators, submitted for publication.
[15] K. Fujimoto, J.M.A. Scherpen, W.S. Gray, Hamiltonian realizations of nonlinear adjoint operators, Proceedings of the IFAC Workshop on Lagrangian and Hamiltonian Methods for Nonlinear Control, Princeton, New Jersey, USA, 2000, pp. 39–44.

[16] K. Fujimoto, J.M.A. Scherpen, W.S. Gray, Hamiltonian realizations of nonlinear adjoint operators, submitted for publication.
[17] M. Furi, M. Martelli, A. Vignoli, Contributions to the spectral theory for nonlinear operators in Banach spaces, Ann. Mat. Pura Appl. 118 (1978) 229–294.
[18] M. Furi, A. Vignoli, A nonlinear spectral approach to surjectivity in Banach spaces, J. Funct. Anal. 20 (1975) 304–318.
[19] W.S. Gray, J.M.A. Scherpen, Hankel operators and Gramians for nonlinear systems, Proceedings of the 37th IEEE Conference on Decision and Control, Tampa, Florida, USA, 1998, pp. 1416–1421.
[20] W.S. Gray, J.M.A. Scherpen, On the nonuniqueness of balanced nonlinear realizations, Proceedings of the 38th IEEE Conference on Decision and Control, Phoenix, Arizona, USA, 1999, pp. 4736–4741.
[21] W.S. Gray, J.M.A. Scherpen, On the nonuniqueness of singular value functions and balanced nonlinear realizations, Systems and Control Letters, 44 (2001) 219–232.
[22] W.S. Gray, J.M.A. Scherpen, Hankel operators, singular value functions and Gramian generalizations for nonlinear systems, submitted for publication.
[23] P. Guillaume, Nonlinear eigenproblems, SIAM J. Matrix Anal. Appl. 20 (3) (1999) 575–595.
[24] S. Jingxian, L. Bendong, Eigenvalues and eigenvectors of nonlinear operators and applications, Nonlinear Anal. Theory Meth. Appl. 29 (11) (1997) 1277–1286.
[25] E.A. Jonckheere, L.M. Silverman, Singular value analysis of deformable systems, Circuits Systems Signal Process. 1 (3–4) (1982) 447–470.
[26] T. Kato, Perturbation Theory for Linear Operators, 2nd Edition, Springer, Heidelberg, 1966.
[27] G.I. Marchuk, V.I. Agoshkov, V.P. Shutyaev, Adjoint Equations and Perturbation Algorithms in Nonlinear Problems, CRC Press, Boca Raton, FL, Inc., 1996.
[28] B.C. Moore, Principal component analysis in linear systems: controllability, observability, and model reduction, IEEE Trans. Automat. Control AC-26 (1) (1981) 17–32.
[29] M.Z. Nashed, Differentiability and related properties of nonlinear operators: Some aspects of the role of differentials in nonlinear functional analysis, in: L.B. Rall (Ed.), Nonlinear Functional Analysis and Applications, Academic Press, New York, 1971, pp. 103–309.
[30] J.W. Neuberger, Existence of a spectrum for nonlinear transformations, Pacific J. Math. 31 (1) (1969) 157–159.
[31] J. Pejsachowicz, A. Vignoli, On differtiability and surjectivity of α-Lipschitz mappings, Ann. Mat. Pura Appl. 101 (1974) 49–63.
[32] A.P. Robertson, W.J. Robertson, Topological Vector Spaces, 2nd Edition, University Press, Cambridge, 1973.
[33] J.M.A. Scherpen, Balancing for nonlinear systems, Systems Control Lett. 21 (1993) 143–153.
[34] J.M.A. Scherpen, Balancing for nonlinear systems, Proceedings of the 1993 European Control Conference 4 (1993) 1838–1843.
[35] J.M.A. Scherpen, W.S. Gray, On singular value functions and Hankel operators for nonlinear systems, Proc. 1999 IEEE American Control Conference, San Diego, California, USA, 1999, pp. 2360–2364.
[36] J.M.A. Scherpen, W.S. Gray, Minimality and local state decompositions of a nonlinear state space realization using energy functions, IEEE Trans. Automat. Control AC-45 (11) (2000) 2079–2086.
[37] J.M.A. Scherpen, A.J. van der Schaft, Normalized coprime factorizations and balancing for unstable nonlinear systems, Int. J. of Contr. 60 (6) (1994) 1193–1222.
[38] V.P. Shutyayev, On calculation of a functional in a nonlinear problem using an adjoint equation, Comput. Math. Math. Phys. 31 (9) (1991) 8–16.
[39] V.S. Sunder, Functional Analysis: Spectral Theory, Birkhäuser, Berlin, 1998.
[40] S. Yamamuro, The adjoints of differentiable mappings, J. Austrailian Math. Soc. 8 (1968) 297–409.
[41] E.H. Zarantonello, The closure of the numerical range contains the spectrum, Pacific J. Math. 22 (3) (1967) 575–595.
[42] V.V. Zharinov, Secondary operational calculus, Proceedings of the International Conference on Secondary Calculus and Cohomological Physics, Moscow, 1997.
[43] K. Zhou, J.C. Doyle, K. Glover, Robust and Optimal Control, Prentice-Hall, Inc., Upper Saddle River, NJ, 1996.